

Finding Influential Institutions in Bibliographic Information Networks

Anubhav Gupta
Department of Computer Science
and Automation
Indian Institute of Science
Bangalore, India
anubhav.gupta@csa.iisc.ernet.in

M. Narasimha Murty
Department of Computer Science
and Automation
Indian Institute of Science
Bangalore, India
mnm@csa.iisc.ernet.in

ABSTRACT

Ranking in bibliographic information networks is a widely studied problem due to its many applications such as advertisement industry, funding, search engines, etc. Most of the existing works on ranking in bibliographic information network are based on ranking of research papers and their authors. But the bibliographic information network can be used for solving other important problems as well. The KDD Cup 2016 competition considers one such problem, which is to measure the impact of research institutions, i.e. to perform ranking of research institutions. The competition took place in three phases. In this paper, we discuss our solutions for ranking institutions in each phase. We participated under team name “anu@TASL” and our solutions achieved the average NDCG@20 score of 0.7483, ranking in eleventh place in the contest.

Keywords

Ranking, Heterogeneous networks, Time-series

1. INTRODUCTION

In recent years, ranking in information network has attracted a lot of attention due to its wide applications in search engines, advertisements, etc. Most of the work on ranking in information networks is focused on analyzing homogeneous network [11, 8], which covers only a small part of the large information network. This information network forms very large heterogeneous networks, that are composed of multiple types of objects connected by links between them. Some of the recent work [16, 13, 10] also takes into consideration, the heterogeneous structure to perform ranking.

Because of the importance associated with the problem of ranking, KDD Cup 2016 is also based on this exciting problem. The task of KDD Cup Data Mining Contest 2016 hosted by Microsoft Azure is finding influential institutions

in the academic network. To be exact, given any upcoming conference such as KDD, the task is to perform ranking of institutions based on predicting how many of their research papers will be accepted in KDD in 2016. The contestants were given the choice to use any publicly available dataset, for predicting next year’s top institutions. We have used Microsoft Academic Graph dataset for this purpose. We have considered this problem as a supervised learning problem and also used temporal information in our ranking model. The competition took place in three phases. In each phase, the contestants were evaluated on one conference which was chosen from some given conferences in that phase. In this paper, we present our solution towards solving this problem. After each phase, we modified our model and used a different algorithm for performing the ranking. In this paper, we describe the algorithms used in each phase individually and demonstrate their performance in all the three phases.

To evaluate the rankings produced by our model, the metric NDCG (Normalized Discounted Cumulative Gain) is used. In information retrieval, Normalized Discounted Cumulative Gain [7] is a standard metric for evaluating rankings. The Discounted Cumulative Gain (DCG) at position n is calculated using the following formula:

$$DCG@n = \sum_{i=1}^n \frac{rel_i}{\log_2(i+1)}$$

Then, NDCG at position n is defined as:

$$NDCG@n = \frac{DCG@n}{IDCG@n}$$

where i is the predicted rank of an institution and rel_i is its true relevance score. For a perfect ranking algorithm, IDCG is equal to DCG producing an NDCG of 1.0. For the calculation of true relevance scores, the number of accepted full research papers from every institution is used.

The rest of the papers is organized as follows. Section 2 describes relevant notation and assumptions made by our approach. Our ranking frameworks are proposed in detail in section 3. In section 4, we perform some experiments and discuss the results of our methods. Finally, we conclude our work in section 5.

2. NOTATION AND ASSUMPTIONS

The competition took place in three phases. In each phase, contestants were asked to predict ranking of institutions for a different set of conferences. Let the number of institutions for which we predict rankings be m . In the competition,

Phase	Conferences
1	SIGIR, SIGMOD, SIGCOMM
2	KDD, ICML
3	FSE, MobiCom, MM

Table 1: List of different conferences for different phases of KDD Cup 2016

$m = 741$. Let us call these institutions I_1, I_2, \dots, I_m . The list of conferences for each of these phases is given in table 1.

Following assumptions were imposed by the KDD Cup organizers for finding the relevance of institutions:

- For any conference, all the accepted papers are equally important. i.e. each accepted paper has equal vote.
- Whenever a research paper is authored by multiple authors, each author is assumed to have made equal contribution to the paper.
- If an author is affiliated with multiple institutions, each institution has equal contribution to the paper.

3. PROPOSED METHODS

In this section, we discuss our approaches in each phase of the competition.

3.1 Solution for Phase-1

In phase-1, contestants were asked to predict institution rankings for SIGIR, SIGMOD and SIGCOMM. In this phase, we tried a simple approach where we predict the ranking score of institution for a conference based on their ranking scores in the same conference in previous years. But here, the ranking scores for previous years are weighted differently for different years. Our intuition is that rank of an institution in a conference in 2016 will depend on its rank in the same conference in year 2015 more than it depends on its rank in year 2014 and so on.

3.1.1 Learning weights of past ranking scores

As discussed above, ranking score (fraction of research papers) of an institution in the conference C_b in year 2016 is obtained by taking the weighted average of ranking scores of this institution in conference C_b in years 2015, 2014, etc. The weights for rankings scores of years 2015, 2014, etc. are learned using Brown's simple exponential smoothing.

Brown's Simple Exponential Smoothing: Brown's simple exponential smoothing (exponentially weighted moving average) [1, 2] is a popular technique for smoothing time series data. This technique has the property that it doesn't treat all the past observations equally and assigns them different weights, such that the most recent observation gets more weight than 2^{nd} most recent and 2^{nd} most recent gets more weight than 3^{rd} most recent, and so on. In the simple exponential smoothing method, forecast for a variable Y at time $t + 1$ is given by

$$Y_{t+1} = \alpha[Y_t + (1 - \alpha)Y_{t-1} + (1 - \alpha)^2Y_{t-2} + \dots]$$

where α is called the *smoothing constant* and takes value in the range $[0, 1]$.

Algorithm 1: RankIns1

Input: Bibliographic Information Network \mathcal{G} and conference C_b to predict rankings for

Result: Rankings of institutions corresponding to the conference C_b

```

1: Initialize  $all\_w \leftarrow [\frac{1}{20}, \frac{2}{20}, \dots, 1]$ 
2: Initialize  $max\_score \leftarrow 0$ 
3: Initialize  $w\_opt \leftarrow 1$ 
4: for  $k = 1$  to  $m$  do
5:    $r_{b,k}^y \leftarrow$  ranking score (fraction of papers) of institute  $I_k$  in year  $y$  with respect to conference  $C_b$ 
6: end for
// Use ranking scores of institutions in year 2015 to learn  $\beta$ 's
7: for  $w$  in  $all\_w$  do
8:   for  $k = 1$  to  $m$  do
9:      $s_{b,k}^{2015} \leftarrow r_{b,k}^{2014} + w \cdot r_{b,k}^{2013} + w^2 \cdot r_{b,k}^{2012} + w^3 \cdot r_{b,k}^{2011}$ 
10:   end for
11:    $score \leftarrow$  NDCG@20 score of ranking  $s_b^{2015}$ , with the true ranking given by  $r_b^{2015}$ 
12:   if  $score \geq max\_score$  then
13:      $max\_score \leftarrow score$ 
14:      $w\_opt \leftarrow w$ 
15:   end if
16: end for
17: for  $k = 1$  to  $m$  do
18:    $s_{b,k}^{2016} \leftarrow r_{b,k}^{2015} + w \cdot r_{b,k}^{2014} + w^2 \cdot r_{b,k}^{2013} + w^3 \cdot r_{b,k}^{2012}$ 
19: end for
20: Normalize  $s_{b,k}^{2016}$  scores so that they sum to 1
21: Output  $s_{b,k}^{2016}$  as relevance scores of institutions

```

In our case, let Y_i 's denote the ranking scores of an institution in different years. We also assume that ranking score of an institution in any year depends only on its ranking scores in previous four years. So, ranking score of an institution in year 2016 is given by

$$Y_{2016} = \alpha[Y_{2015} + (1 - \alpha)Y_{2014} + (1 - \alpha)^2Y_{2013} + (1 - \alpha)^3Y_{2012}]$$

Now, since α is same for all institutions given a conference, for the purpose of ranking, we only need to learn the parameter β such that

$$Y_{2016} = Y_{2015} + \beta Y_{2014} + \beta^2 Y_{2013} + \beta^3 Y_{2012}$$

To learn β for a conference C_b , we use the data upto year 2014 for training and find the parameter β that optimizes the NDCG@20 score for ranking of institutions for conference C_b in year 2015. Then, this β is used to predict the ranking of institutions for year 2016.

The algorithm for ranking institutions is summarized in Algorithm 1

3.2 Solution for Phase-2

In phase-2, contestants were asked to predict rankings for the conferences KDD and ICML. The algorithms used for ranking institutions in this phase is same as the one used in phase-1, which is given by Algorithm 1.

3.3 Solution for Phase-3

In the third phase, contestants were asked to predict ranking of institutions for the conferences FSE, MobiCom and MM. In this section, we discuss the algorithm used in the

phase-3 of the competition – **RankIns3**. This algorithm is different from the algorithm used in previous two phases. We are given the name of a conference, say C_b , as a query, and our task is to predict the ranking of institutions based on predicting how many of their research papers will be accepted in this conference in year 2016. For this purpose, we have used the information provided by heterogeneous bibliographic network upto year 2015 for training the ranking model.

3.3.1 Intuition Behind our Approach

Our ranking model is based on the intuition that rank of the institution I_k for a conference C_b depends on ranks of other similar institutions in conference C_b as well as on its rank in conferences similar to C_b . Also, rank of I_k in conference C_b depends on rank of I_k in the past instances of C_b , i.e. rank of I_k in conference C_b in 2015, 2014, etc. Therefore, conferences and institutions are represented as feature vectors as such a representation allows the similarity between institutions and conferences, and between the institutions to be measured in terms of cosine of their feature vectors.

The RankIns2 algorithm works as follows. The existing problem is first transformed into a learning-to-rank problem [9] and then learning-to-rank techniques such as RankSVM [5], RankBoost [4], AdaRank [14] can directly be used to solve the learning problem. To rank institutions, RankIns2 first constructs the data matrix for year 2016 and the data matrices upto year 2015 are used to train the learning-to-rank model. Then this learned model is used to predict ranks of institutions for data matrix of year 2016.

3.3.2 Construction of Features Vectors

RankIns2 uses the following method in order to construct feature vectors of institutions:

- Authors and keywords associated with papers are used to identify the institutions.
- Instead of using individual authors, clusters of authors are used as features. To perform the clustering of authors, methods such as BAGC [15] and SI-Cluster [17] can be used. Let the number of clusters be K . Let these K clusters be represented by their respective centers a_1, a_2, \dots, a_K .
- For the datasets in which topic information is not available, methods such as Latent Dirichlet Allocation [3] and Probabilistic Latent Semantic Indexing [6] can be used to find topics of research papers.

Feature vectors for all the conferences are constructed in the same way. Let us say, feature vector of institution I_k is given by:

$$I_k = \overbrace{[\alpha_{i_1} \ \alpha_{i_2} \ \dots \ \alpha_{i_K}]}^{\text{Author Clusters}} \overbrace{[\beta_{i_1} \ \beta_{i_2} \ \dots \ \beta_{i_s}]}^{\text{Topics}}$$

and feature vector of conference C_b is given by:

$$C_b = \overbrace{[\alpha_{c_1} \ \alpha_{c_2} \ \dots \ \alpha_{c_K}]}^{\text{Author Clusters}} \overbrace{[\beta_{c_1} \ \beta_{c_2} \ \dots \ \beta_{c_s}]}^{\text{Topics}}$$

Here, s is the number of distinct topics in the dataset. Now, the feature vector of institution I_k corresponding to

conference C_b is obtained by taking the element-wise multiplication of their feature vectors, and is given by

$$I_k^{(b)} = I_k \odot C_b$$

$$I_k^{(b)} = [\alpha_{i_1} \alpha_{c_1} \ \dots \ \alpha_{i_K} \alpha_{c_K} \ \beta_{i_1} \beta_{c_1} \ \dots \ \beta_{i_s} \beta_{c_s}]$$

Let these feature vectors corresponding to conference C_b be together denoted by matrix M_b , which is defined as follows:

$$M_b = \begin{pmatrix} \overleftarrow{\hspace{1cm}} I_1^{(b)} \overrightarrow{\hspace{1cm}} \\ \overleftarrow{\hspace{1cm}} I_2^{(b)} \overrightarrow{\hspace{1cm}} \\ \vdots \\ \overleftarrow{\hspace{1cm}} I_k^{(b)} \overrightarrow{\hspace{1cm}} \end{pmatrix}_{k \times (K+s)}$$

So, the matrix M_b is a $k \times d$ matrix, where $d = K + s$.

3.3.3 Constructing Data Matrix for Year 2016

Here, we describe the procedure to find data matrix for year 2016 corresponding to conference C_b . Since, the institutions in data matrices are represented in the form of clusters of authors and topics, the data matrix for year 2016 can be obtained by using data matrices of past years because any new research paper will belong to either some of the existing topics or some mixture of them, and will be written by an author belonging to one of the ranked categories. Let us assume that data matrix of any year depends only on the data matrices of previous three years. So, the data matrix of 2016 depends only on data matrices of years 2015, 2014 and 2013. Let us denote by $M_b^{(y)}$, data matrix for year y corresponding to conference C_b .

To learn the data matrix for year 2016, we make the assumption that data matrix of any year y is a linear combination of data matrices of years $y - 1$, $y - 2$ and $y - 3$. i.e. data matrix of year 2016

$$\hat{M}_b^{(2016)} = w_1 \cdot M_b^{2015} + w_2 \cdot M_b^{2014} + w_3 \cdot M_b^{2013} \quad (1)$$

where w_1 , w_2 and w_3 are the weights of data matrices for years 2015, 2014 and 2013 respectively. Let us represent these weights by a 3-dimensional vector $\mathbf{w} = [w_1 \ w_2 \ w_3]^T$. To learn these weights, we first learn the initial weight vector $\mathbf{w}^{(0)}$ by solving the following optimization problem using the data upto year 2015 for training:

$$\mathbf{w}^{(0)} = \arg \min_{\mathbf{w}} \left\| M_b^{2015} - \hat{M}_b^{2015} \right\|_F^2 \quad (2)$$

where

$$\hat{M}_b^{2015} = w_1 \cdot M_b^{2014} + w_2 \cdot M_b^{2013} + w_3 \cdot M_b^{2012}$$

To find the solution of the above equation, we introduce some additional notation. Let n be the number of authors in the dataset. Let us denote by R , a $(n \times d \times 3)$ matrix tensor, such that

$$R_{ij:} = [(M_b^{2014})_{ij} \ (M_b^{2013})_{ij} \ (M_b^{2012})_{ij}]$$

Let us define a matrix $X \in \mathbb{R}^{(md) \times 3}$ whose rows are the tube fibers of the tensor R , as follows:

$$X = [R_{11:} \ R_{12:} \ \dots \ R_{md:}]^T$$

Also, let

$$\mathbf{z} = [(M_b^{2015})_{11} \ (M_b^{2015})_{12} \ \dots \ (M_b^{2015})_{md}]^T$$

Then, one can easily see following relationship:

$$\left\| M_b^{2015} - \hat{M}_b^{2015} \right\|_F^2 = \|\mathbf{z} - X\mathbf{w}\|^2$$

Using this equality, the problem given in equation (2) reduces to

$$\mathbf{w}^{(0)} = \arg \min_{\mathbf{w}} \|\mathbf{z} - X\mathbf{w}\|^2$$

The above equation has a closed form solution, which is obtained by equating the gradient of the above equation to zero. The solution is given by:

$$\mathbf{w}^{(0)} = (X^T X)^{-1} X^T \mathbf{z} \quad (3)$$

After learning the initial weight vector $\mathbf{w}^{(0)}$, the algorithm iteratively updates this weight vector by using the data of previous years. We use data upto years 2014, 2013, ... at each iteration of the algorithm. At l^{th} iteration, the following optimization problem is solved to update the weight vector $\mathbf{w} = [w_1 \ w_2 \ w_3]^T$:

$$\begin{aligned} \mathbf{w}^{(l)} = \arg \min_{\mathbf{w}} & \left\| M_b^{(2015-l)} - \hat{M}_b^{(2015-l)} \right\|_F^2 \\ & + \lambda_l \left\| \mathbf{w}^{(l-1)} - \mathbf{w} \right\|^2 \end{aligned} \quad (4)$$

Here,

$$\hat{M}_b^{(2015-l)} = w_1 \cdot M_b^{(2014-l)} + w_2 \cdot M_b^{(2013-l)} + w_3 \cdot M_b^{(2012-l)}$$

The λ_l 's above are the hyper-parameters, used to make sure that recent links in the network are given more importance than past links. The second term in the above equation is a regularizer and it ensures that the updated weight vector is not too far away from its previous value.

Again, for easy calculation, we introduce some notation. Let us denote by $S^{(l)}$, a $(n \times d \times 3)$ matrix tensor, such that

$$S_{ij:}^{(l)} = [(M_b^{(2014-l)})_{ij} \ (M_b^{(2013-l)})_{ij} \ (M_b^{(2012-l)})_{ij}]$$

Let us define the matrix $X_l \in \mathbb{R}^{(md) \times 3}$ whose rows are the tube fibers of matrix tensor $S^{(l)}$, as follows:

$$X_l = [S_{11:}^{(l)} \ S_{12:}^{(l)} \ \dots \ S_{md:}^{(l)}]^T$$

Also, let us define the vector \mathbf{z}_l as follows:

$$\mathbf{z}_l = [(M_b^{(2015-l)})_{11} \ (M_b^{(2015-l)})_{12} \ \dots \ (M_b^{(2015-l)})_{md}]^T$$

Then, one can observe following relationship:

$$\left\| M_b^{(2015-l)} - \hat{M}_b^{(2015-l)} \right\|_F^2 = \|\mathbf{z}_l - X_l \mathbf{w}\|^2$$

Using the above equality, the problem given by equation (4) reduces to

$$\mathbf{w}^{(l)} = \arg \min_{\mathbf{w}} \|\mathbf{z}_l - X_l \mathbf{w}\|^2 + \lambda_l \left\| \mathbf{w}^{(l-1)} - \mathbf{w} \right\|^2$$

By taking the gradient of above equation and equating it to zero, we get:

$$2X^T(X_l \mathbf{w} - \mathbf{z}_l) - 2\lambda_l(\mathbf{w}^{(l-1)} - \mathbf{w}) = \mathbf{0}$$

Algorithm 2: RankIns2

Input: Bibliographic Information Network \mathcal{G} and conference C_b to predict rankings for

Result: Rankings of institutions corresponding to conference C_b

- 1: Construct feature vectors for all the institutions and all the conferences
- 2: Construct feature vectors corresponding to every institution and conference pair
- 3: Create matrices $M_b^{(i)}$ for $i = 2011, 2012, \dots, 2015$ for conference C_b
- 4: Create matrix X and vector \mathbf{z} as discussed in Section 3.3.3
- 5: Initialize

$$\mathbf{w}^{(0)} \leftarrow (X^T X)^{-1} X^T \mathbf{z}$$

6: **for** $l = 1$ to u **do**

7: Create matrix X_l and vector \mathbf{z}_l as discussed in Section 3.3.3

8: Update the weight vector \mathbf{w} as follows:

$$\mathbf{w}^{(l)} \leftarrow (X_l^T X_l + \lambda_l I)^{-1} (X_l^T \mathbf{z}_l + \lambda_l \mathbf{w}^{(0)})$$

9: **end for**

10: Final weight vector $\mathbf{w}^* \leftarrow \mathbf{w}^{(u)}$

11: Construct data matrix for year y as

$$\hat{M}_b^{2016} = \mathbf{w}_1^* \cdot M_b^{2015} + \mathbf{w}_2^* \cdot M_b^{2014} + \mathbf{w}_3^* \cdot M_b^{2013}$$

12: Train the learning model with data upto year 2015

13: Use this model to make predictions for all the institutions in \hat{M}_b^{2016}

By, solving the above equation, we obtain the following solution for $\mathbf{w}^{(l)}$

$$\mathbf{w}^{(l)} = (X_l^T X_l + \lambda_l I)^{-1} (X_l^T \mathbf{z}_l + \lambda_l \mathbf{w}^{(l-1)}) \quad (5)$$

3.3.4 Solving the learning problem

Using the weights learned above, we create the data matrix of all institutions for year 2016 corresponding to conference C_b as given in equation (1). Then, using the data matrices upto year 2015 for all the conferences for training, we treat the problem as a regression problem and predict the scores for institutions using random forest regression method. These scores are the predicted relevance scores of institutions.

Algorithm 2 summarizes the working of our method. The input to RankIns2 is a bibliographic information network \mathcal{G} and a conference C_b to predict rankings for. RankIns2 then predicts the rankings of institutions corresponding to this given conference.

4. EXPERIMENTS AND RESULTS

4.1 Dataset Description

For performing the ranking of institutions, we have used the Microsoft Academic Graph (MAG) dataset which is a huge dataset containing information about scientific documents from different research domains. Figure 2 shows the schema of MAG.

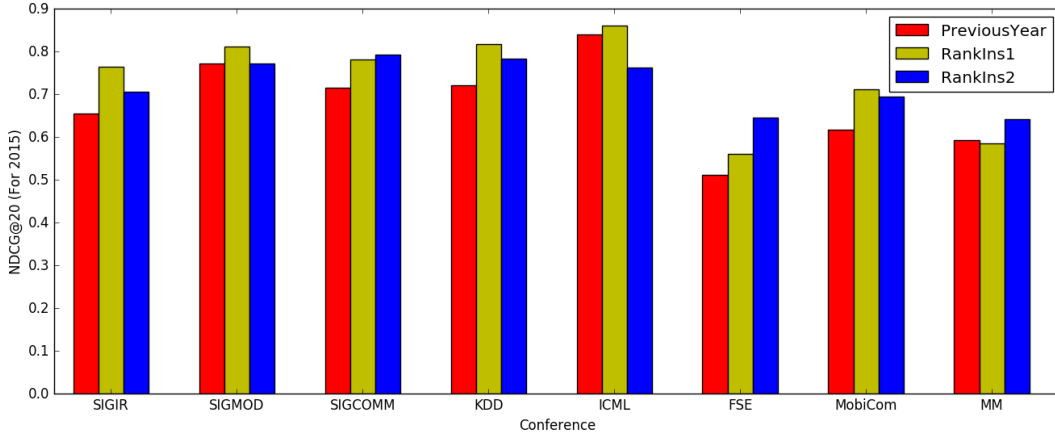


Figure 1: Performance of RankIns1 and RankIns2 on validation dataset

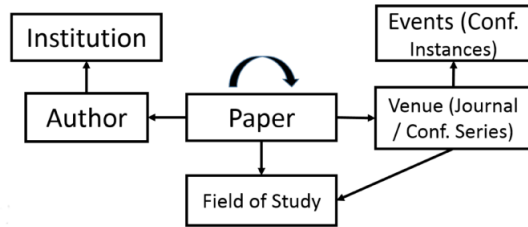


Figure 2: Schema of Microsoft academic graph dataset, adapted from [12]

Along with the MAG dataset, organizers of KDD Cup additionally provided two separate files for the purpose of this competition. These two files contain the names of institutions for which we have to predict the rankings and details of full research papers accepted in the asked conferences during the time period 2011 - 2015.

We merged MAG dataset with the given two files and then divided the dataset on a yearwise basis. i.e. we created several small heterogeneous bibliographic networks by considering the research papers of year 2011, 2012, ..., 2015 respectively.

4.2 Evaluation

To evaluate the proposed ranking algorithms, the metric *Normalized Discounted Cumulative Gain* is used:

Normalized Discounted Cumulative Gain: In information retrieval, Normalized Discounted Cumulative Gain (NDCG) [7] is a standard metric for evaluating rankings. The Discounted Cumulative Gain (DCG) at position n is calculated using the following formula:

$$DCG@n = \sum_{i=1}^n \frac{rel_i}{\log_2(i+1)}$$

Then, NDCG at position n is defined as:

$$NDCG@n = \frac{DCG@n}{IDCG@n}$$

Phase	Conference	NDCG@20	Algorithm
1	SIGIR	0.7364464224	RankIns1
2	KDD	0.7815387475	RankIns1
3	MM	0.7210261755	RankIns2

Table 2: Performance of RankIns1 and RankIns2 on test dataset

where i is the predicted rank of an institution and rel_i is its true relevance score. For a perfect ranking algorithm, IDCG is equal to DCG producing an NDCG of 1.0.

4.3 Results and Discussion

To evaluate the proposed methods, one conference out of all the given conferences was selected in each phase, to form the test dataset. Name of the conference selected and the results of our methods for these conferences are given in Table 2. We also create a validation set, where we predicted the ranks of institutions for all the conferences appearing in the three phases in year 2015. The performance of RankIns1 and RankIns2 for each of these conferences is given in figure 1. To compare our methods, we also take one more ranking of institutions. This ranking, called **PreviousYear**, ranks an institution corresponding to a conference in year y , same as its rank in $y - 1$ for this conference.

In our implementation of RankIns2, we have used $u = 2$ and $\lambda_l = 200$. The number of clusters of authors used to create feature vectors is 500.

The results on validation set show that proposed methods RankIns1 and RankIns2 outperforms PreviousYear for 7 out of 8 conferences. From results, however, it is not clear whether RankIns1 should be preferred over RankIns2 or not. But RankIns1 has a clear drawback compared to RankIns2. For the conferences where there is a sudden change in the rankings of institutions in the current year, but rankings are similar in previous years, rankings predicted by RankIns1 will be heavily affected by current year rankings (which seems to be an outlier).

Data preprocessing was performed on an Intel Xeon E5-

2640 v3 2.60 GHz (Haswell-based) machine with 128 GB of memory. All the experiments were performed on an Intel Core I5 4200U machine with 4GB of memory. Our codes are written in Python 2.7.11. Our codes are all single-threaded.

5. CONCLUSION

In this paper, we have discussed the problem of ranking institutions and proposed our methods RankIns1 and RankIns2. Both RankIns1 and RankIns2 consider the ranking problem as a supervised learning problem and both the approaches consider the bibliographic data in the form of time-series. RankIns1 predicts the ranking scores of institutions by giving weights to their ranking scores in previous years. Whereas, RankIns2 works by transforming the problem into a learning-to-rank problem and then uses the learning-to-rank framework to predict the rankings of institutions.

6. ACKNOWLEDGMENTS

We would like to acknowledge SIGKDD for organizing such a challenging and inspiring competition. We also thank Microsoft for providing the dataset and the user friendly platform for the competition. Finally, we would like to acknowledge Sharad Nandanwar of IISc for the fruitful discussions we had with him.

7. REFERENCES

- [1] <http://people.duke.edu/~rnau/411avg.htm>.
- [2] https://en.wikipedia.org/wiki/exponential_smoothing.
- [3] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.
- [4] Y. Freund, R. Iyer, R. E. Schapire, and Y. Singer. An efficient boosting algorithm for combining preferences. *J. Mach. Learn. Res.*, 4:933–969, 2003.
- [5] R. Herbrich, T. Graepel, and K. Obermayer. *Large margin rank boundaries for ordinal regression*. MIT Press, Cambridge, MA, 2000.
- [6] T. Hofmann. Probabilistic latent semantic indexing. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, pages 50–57. ACM, 1999.
- [7] K. Järvelin and J. Kekäläinen. Ir evaluation methods for retrieving highly relevant documents. In *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*, pages 41–48. ACM, 2000.
- [8] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5):604–632, 1999.
- [9] T.-Y. Liu. Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval*, 3(3):225–331, 2009.
- [10] Q. Meng and P. J. Kennedy. Discovering influential authors in heterogeneous academic networks by a co-ranking method. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*, pages 1029–1036. ACM, 2013.
- [11] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: bringing order to the web. Stanford InfoLab, 1999.
- [12] A. Sinha, Z. Shen, Y. Song, H. Ma, D. Eide, B.-j. P. Hsu, and K. Wang. An overview of microsoft academic service (mas) and applications. In *Proceedings of the 24th International Conference on World Wide Web Companion*, pages 243–246. International World Wide Web Conferences Steering Committee, 2015.
- [13] Y. Sun, J. Han, P. Zhao, Z. Yin, H. Cheng, and T. Wu. Rankclus: integrating clustering with ranking for heterogeneous information network analysis. In *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology*, pages 565–576. ACM, 2009.
- [14] J. Xu and H. Li. Adarank: a boosting algorithm for information retrieval. In *SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 391–398, New York, NY, USA, 2007. ACM.
- [15] Z. Xu, Y. Ke, Y. Wang, H. Cheng, and J. Cheng. A model-based approach to attributed graph clustering. In *Proceedings of the 2012 ACM SIGMOD international conference on management of data*, pages 505–516. ACM, 2012.
- [16] D. Zhou, S. A. Orshanskiy, H. Zha, and C. L. Giles. Co-ranking authors and documents in a heterogeneous network. In *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*, pages 739–744. IEEE, 2007.
- [17] Y. Zhou and L. Liu. Social influence based clustering of heterogeneous information networks. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 338–346. ACM, 2013.